Is Swiss Telecommunications a Natural Monopoly? An Evaluation of Empirical Evidence

S. Buehler*

This Draft: April 19, 2000

ABSTRACT: Based upon time series data published by PTT prior to regulatory reform, this paper investigates whether Swiss telecommunications qualifies as a natural monopoly. Employing the subadditivity concept for multiproduct industries, alternative specifications of quadratic cost functions are estimated. The results of these estimations are ambiguous and demonstrate the difficulty of empirically substantiating a natural monopoly claim. It is argued that the natural monopoly concept is of limited usefulness for informing policy decisions.

Keywords: telecommunications, natural monopoly, time series.

JEL: C32, L 43, L96.

---

*Socioeconomic Institute, University of Zurich, Hottingerstr. 10, 8032 Zurich, Switzerland; phone: +41-1-634 06 06; fax: +41-1-634 49 87;
e-mail: sbuehler@sozoec.unizh.ch.

Computations were conducted using Autobox, Eviews 2.0 and Matlab 5. I am grateful to Klaus Edel, Armin Schmutzler, Winfried Stier, Jörg Wild and Marc Wildi for helpful comments and suggestions. Financial support by the Swiss National Science Foundation is gratefully acknowledged.

1 Introduction

It is fair to say that only a few years ago, there was a broad consensus that the telecommunications industry is a natural monopoly due to large economies of scale and scope in the provision of telecommunications services. Competition in the provision of services—not to mention infrastructure—was therefore deemed to be out of the question. Recently, however, this view has been challenged. Technological progress has altered the industry specific cost conditions by driving down the cost of adding another participant to a given network (i.e. providing the "local loop") and making this cost depend less on the number of existing participants, thereby reducing the importance of economies of scale (see Bond 1997, pp. 11). This process has thoroughly changed the telecommunications industry in most OECD countries and appears to have had its biggest impact in the mobile sector, where network operators began to compete in prices and quality for potential costumers earlier than elsewhere. In the case of Switzerland for example, the former monopoly operator Swisscom is facing both service and infrastructure competition from two other licence holders (Diax and Orange) since the enacting of the new Telecommunications Act (Fernmeldegesetz, FMG) and the accessory ordinances in January 1998. Technological progress has also generated opportunities for firms operating in related network industries, such as electricity and television, to enter telecommunications. By adapting their infrastructure facilities to two-way communication, they are (at least potentially) able to implement effective intermodal competition (see Bräutigam 1979). The increasing proximity of these industries is often called "convergence" and has recently received much attention both in the telecommunications policy literature and in the general public.

Closely related to the conviction that telecommunications is a natural monopoly was the claim that the industry is plagued by systematic market failure. Recent economic research has demonstrated, however, that there is a large set of regulatory means available to cope with market failure in network industries. Important elements of these regulatory means are

- appropriate rules for "interconnection" which ensure non-discriminating network access and interoperability between competing networks,[1] and

- universal service prescriptions that do no distort economic decisions of consumers or producers.[2]

In Switzerland, the revised FMG has brought about a regulatory reform which aims at implementing effective competition in the telecommunications industry. It entails both interconnection regulations and universal service prescriptions and is in many ways fairly similar the regulatory frameworks adopted in the member states of the European union.[3]

But what if Swiss telecommunications remains a natural monopoly, e.g. due to the geographic particularities? Does the implementation of competition which is partially brought about by asymmetric regulatory treatment of the incumbent and potential competitors[4] destroy an efficient industry structure? And can one determine whether the introduction of competition is desirable from a normative point of view? These questions have played an important role in the privatization and divesture process both in the US and Great Britain for years. It is therefore not surprising that the economic literature provides a wealth of empirical studies concerning these issues.[5] Interestingly, the results of these studies are rather inconsistent, some supporting the natural monopoly claim, many rejecting it, depending on the methods and data used (see Berg and Tschirhart 1995, 104).[6]

In view of these facts, it is intriguing that for Switzerland there is no empirical evidence whatsoever on the natural monopoly question.[7] So far,

[1]See Mitchell et al. (1995) and Laffont and Tirole (1996) for non-technical surveys of the relevant issues in "one-way" interconnection. See Buehler (1999) for an analysis of "two-way" interconnection.

[2]See Noam (1994) and Blackman (1995) and the literature cited therein for further details.

[3]See Buehler (1998) for a survey of the regulatory reform in Switzerland.

[4]See Knieps (1998).

[5]See Panzar (1989) for a survey of early contributions as well as Shin and Ying (1992, p. 172) and Berg and Tschirhart (1995) and the references therein for more recent work.

[6]For example, Evans and Heckman (1983) reject the natural monopoly hypothesis for AT&T. Charnes et al. (1988), however, supposedly using the same data base and the same functional form as Evans and Heckman, reach the opposite conclusion by applying goal programming rather than regression analysis. Evans and Heckman (1988) in turn qualify the comparison as impractical due to differences in the data and the model.

[7]There is a related paper by Manzini and Thalman (1994) which analyzes the efficiency

the market reform in Switzerland has therefore been conducted with very little information about the cost structure in the industry. In this paper, we attempt to fill part of this gap by

(i) estimating the industry's cost function prior to regulatory reform,

(ii) determining whether the estimated cost function exhibits subadditivity or not and

(iii) discussing the policy conclusions that can be drawn from this analysis.

The study is based on the subadditivity concept for multiproduct industries which defines the necessary conditions for the existence of a natural monopoly (see Jamison 1998). Due to the fact that during the time of observation PTT was operating as a monopolist, a simple time series approach as proposed by Hunt and Lynk (1990) is employed.

The plan of the paper is as follows. Section 2 reviews the concept of natural monopoly applied in this study, describes the requirements for a proper empirical cost function and explains the reasons for choosing a quadratic as opposed to a Cobb-Douglas or a translog functional form. Section 3 surveys the available data, explains the methods employed to transform it accordingly and discusses the problems involved in this process. Section 4 presents results from these estimations and alternative strategies to refine them. Section 5 discusses policy implications.

2 The Basic Set-Up

To start with, we briefly review the relevant concepts of natural monopoly and subadditivity used in this study and then proceed to define what characteristics a proper empirical cost function has to satisfy.

---

of PTT at producing telecommunications services. However, as will be pointed out in section 2.2, the Cobb-Douglas approach employed in their work cannot be used for testing for a natural monopoly.

2.1 Preliminaries

The 'canonical' definition of a natural monopoly is as follows:

Definition 1 (*Baumol et al.* (*1982, 17*)): *An industry is said to be a* ***natural monopoly*** *if, over the entire relevant range of outputs, the firms' cost function is subadditive.*

Consequently, we need to define under what conditions a cost function is subadditive. Roughly speaking, subadditivity means that the production of a given output vector[8] by a single firm is less costly than by any combination of smaller firms. More formally, the following definition holds:

Definition 2 (*Baumol et al.* (*1982, 17*)): *A cost function* $C(\cdot)$ *is globally* ***subadditive*** *if and only if for any output combination* $\mathbf{y}^1, ..., \mathbf{y}^m$, $\sum_{i=1}^m \mathbf{y}^i = \mathbf{y}$ *and* $\mathbf{y}^i \neq \mathbf{y}$

$$C(\mathbf{y}) < \sum_{i=1}^{m} C(\mathbf{y}^i), \tag{1}$$

where $\mathbf{y}^i = [y_1^i, ..., y_n^i]$ *denotes the output vector of firm* $i = 1, ..., m$.

This concept is not easily applied as it requires information about the stand-alone cost of an individual output's production which is usually not available. However, Baumol et al. (1982, p. 187) prove that the simultaneous existence of strictly *decreasing ray average costs*[9] and *transray convexity*[10]

[8]Note that subadditivity thus implicitly refers to a given set of products. This restriction makes it an instrument hard to handle if technological progress allows the penetration of a market by new products (see section 4 for further details).

[9]Decreasing ray average costs can be interpreted to be the multiproduct equivalent to economics of scale and are defined as follows (Baumol 1977, p. 810)

$$\frac{C(vy_1, ..., vy_m)}{v} < \frac{C(wy_1, ..., wy_m)}{w}, \qquad \forall\ v > w,$$

where v and w measure the distance on a ray through $\mathbf{y}^i = \left[y_1^i, ..., y_n^i\right]$.

[10]A cost function is called transray convex if there is a set of positive constants $k_1, ..., k_n$ such that for two output vectors $\mathbf{y}^a = [y_1^a, ..., y_n^a]$ and $\mathbf{y}^b = \left[y_1^b, ..., y_n^b\right]$ and with $\sum_{i=1}^n k_i y_i^a = \sum_{i=1}^n k_i y_i^b = \sum_{i=1}^n k_i y_i$ the following holds (Baumol 1977, 811):

$$C(\theta\mathbf{y}^a + (1-\theta)\mathbf{y}^b) \leq \theta C(\mathbf{y}^a) + (1-\theta)C(\mathbf{y}^b), \qquad \theta \in [0,1].$$

ensures *local* subadditivity of the cost function (around an observable output $\mathbf{y}$). For a given cost function, the existence of these characteristics is much more easily tested than global subadditivity, hence we apply the local concept in the remainder of this paper.

2.2 Proper Empirical Cost Functions

We now consider the type of function that should be used to estimate costs in the telecommunications industry. Any function used to estimate costs has to be a "proper" cost function (Röller 1990a, 203) measuring minimum cost for a given output $\mathbf{y}$ and fixed input prices $\mathbf{w}$. We shall further require that the function

(i) is nondecreasing in $\mathbf{y}$, i.e $\partial C(\mathbf{y}, \mathbf{w})/\partial y_k \geq 0, \forall k = 1, ..., m$;

(ii) allows easy handling of output vectors with zero elements;

(iii) does not anticipate the results of the estimation ("substantial flexibility"), and

(iv) is empirically practical.

Even though these requirements seem rather obvious, they provide some guidance for choosing the appropriate type of function. For example, both requirement (ii) and (iii) exclude the use of a Cobb-Douglas function. This is easily seen for requirement (ii): if any single output is zero, aggregate Cobb-Douglas cost has to be zero as well. More importantly, with respect to requirement (iii), the following argument applies: by differentiating $C(\mathbf{y}, \mathbf{w}) = \alpha \prod_{k=1}^{n} y_k^{\beta_k}$ we obtain $\partial^2 C(\cdot)/\partial y_i \partial y_j = \beta_i \beta_j^2 C(\cdot)/y_i y_j > 0$, for positive marginal costs and outputs. Hence, the Cobb-Douglas function does not allow for transray convexity (i.e. $\partial^2 C(\cdot)/\partial y_i \partial y_j < 0$) by definition.

Based on the more recent empirical literature, the most obvious choice would be a translog cost function similar to those used in the influential studies by Evans and Heckman (1984), Shing and Ying (1992) and numerous others.[11] Yet, given the requirements (ii) to (iv) mentioned above, there are three arguments against it: First, the translog approach does not allow easy handling of zero elements in output vectors in general. While a Box-Cox transformation of the form $y_k^{(\lambda)} = \left(y_k^\lambda - 1\right)/\lambda$ of the data (with $\lambda > 0$) can help to circumvent that problem, it also requires the estimation of the additional parameter λ which might turn out to be negative. In that case, however, the Box-Cox transformation is not defined (see Greene 1993, 334). Second, the "substantial flexibility" of a translog function appears to be excessively high due to the so-called "flip-flop"-property. Loosely speaking, this means that when estimating a translog function, even a marginal change of parameters close to zero might be enough to completely overturn the shape of the estimated cost plane and therefore the conclusions to be drawn from the analysis (see Röller 1990b, pp. 1665). Third, the function ought to have a minimum of parameters to estimate. A comparison of the translog (see fn. 11) and the quadratic cost function

$$C(\mathbf{y}, \mathbf{w}) = F + \sum_k \beta_k y_k + \frac{1}{2} \sum_k \sum_l \beta_{kl} y_k y_l \tag{2}$$

proposed by Baumol and Braunstein (1977), Baumol et al. (1982) and Röller (1990b) establishes the clear advantages of the latter with respect to simplicity.[12]

Thus, in the following we shall use the quadratic cost function defined in (2) with $F \geq 0$ and $\beta_{kl} = \beta_{lk}$. According to this specification, aggregate cost $C(\cdot)$ depends on fixed cost F and the linear and quadratic output terms in y_k and y_l. F, β_k and β_{kl} are assumed to be functions of the non-observable

[11] The translog function is typically specified as (see Röller 1990b, 1663)

$$\begin{aligned} \log C(\cdot) &= \alpha_0 + \sum_{k=1}^{n+1} \alpha_k \ln(y_k) + \sum_{k=1}^{m} \beta_k \ln(p_k) + \frac{1}{2} \sum_{k=1}^{n+1} \sum_{l=1}^{n+1} \alpha_{kl} \ln(y_k) \ln(y_l) \\ &\quad + \frac{1}{2} \sum_{k=1}^{m} \sum_{l=1}^{m} \beta_{kl} \ln(p_k) \ln(p_l) + \sum_{k=1}^{n+1} \sum_{l=1}^{k} \delta_{kl} \ln(y_k) \ln(p_l). \end{aligned}$$

with p_k and p_l denoting input prices and $y_{n+1} = t$ technological progress.

[12] In fact, the quadratic functional form is a restricted version of the translog when the data is in logarithmic form.

input prices $\mathbf{w}$.[13]

3 Data

The most important data source for this study are the statistical yearbooks of PTT from 1963 to 1992 (30 years). This is the time period for which largely consistent data for local, long distance and international calls as well as the cost of providing these services is available.[14] These time series are used to estimate a quadratic cost function $C(\mathbf{y}, \mathbf{w})$. *Local*, *long distance* and *international calls* are different outputs and thus elements of the output vector $\mathbf{y}$. C is the observed cost of providing telephony services and $\mathbf{w}$ denotes the (non-observed) vector of input prices.[15]

To further simplify and be more specific, let y_{1t} denote the logarithm of local calls at time t and y_{2t} the log of the combined outputs of long distance and international calls, i.e. $\mathbf{y} = [y_{1t}, y_{2t}]$. Let c_t be the log of costs deflated by the consumer price index of the Swiss federal statistical office (BfS), for want of a more adequate input price index. Note that due to varying data formats all series are normalized at one for the first observation, hence after taking logs all series start at zero.[16] The quadratic term of the cost function is therefore $x_t = y_{1t}y_{2t}$. Figure 1 plots the data set.

The data clearly demonstrate that the production of telecommunications services has strongly increased during the time of observation, indicating the increasing demand for telephony over time. For trend-stationary variables, the correct procedure to achieve stationarity is detrending. However, calculations have shown that due to the small variance of the variables, detrending results in an almost complete loss of the information on the relation between

[13]It is easily shown that the quadratic cost function actually fulfills the requirements above.

[14]Note that calls free of charge are not considered as they are probably not subject to cost minimizing behavior.

[15]The cost figures are taken from columns 15 and 35 ("telephone") of the cost accounting in the statistical yearbooks of PTT.

[16]In the statistical yearbook, outputs are measured either in number of calls or number of minutes charged. For the purposes of this study, the variables are expressed in logarithmic form since it is unlikely that linear combinations of the levels of variables could be cointegrated (see Hunt and Lynk 1990, 234 and Banerjee et al. 1993, 192 ff.).

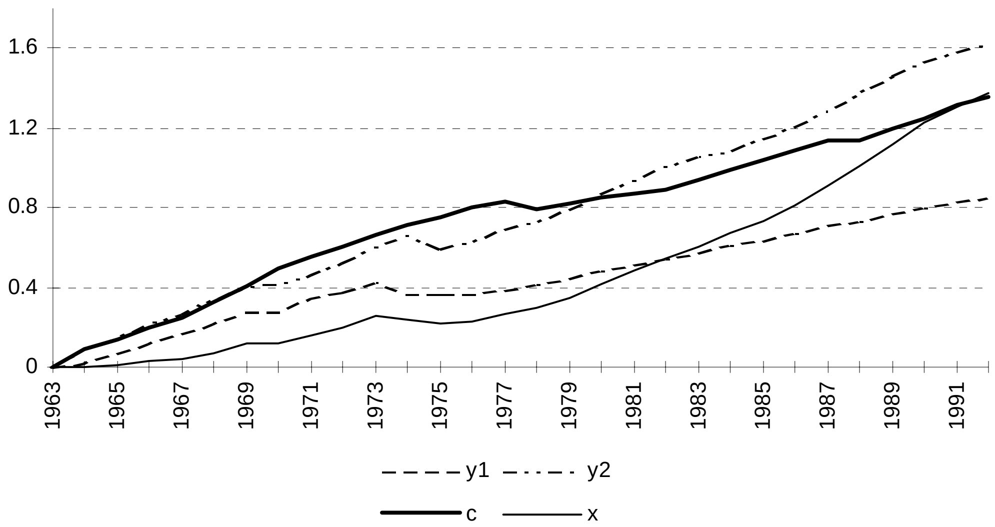

Figure 1: The data set [Source: PTT (several years); BfS].

the variables and hence of the information we are ultimately interested in.[17] In the case of unit root processes, detrending is not appropriate since it does not remove the non-stationarity (Hamilton 1994, 444), and the regression of a unit root process on a constant and a linear trend is itself "spurious" (see Banerjee et al. 1993, 191). Instead of detrending we therefore analyze the order of integration of each variable in the sample (see Figure 2 and Table 1).

Inspection of Figure 2 indicates that not only the levels but also the first differences of the series are non-stationary. This visual impression needs to be supplemented by more formal statistical unit root testing. There is a large number of tests for determining the number of unit roots with different properties regarding the power against the alternative hypothesis. We will follow the dominant procedure in applied work and adopt the consistent sequential *Dickey-Pantula procedure*, which suggests to start testing with the highest possible order of integration and then testing down the integration

[17]This is true both for a simple linear trend and the well-known Hodrick-Prescott (high-pass) filter with the smoothing parameter $\eta = 100$ for yearly data.

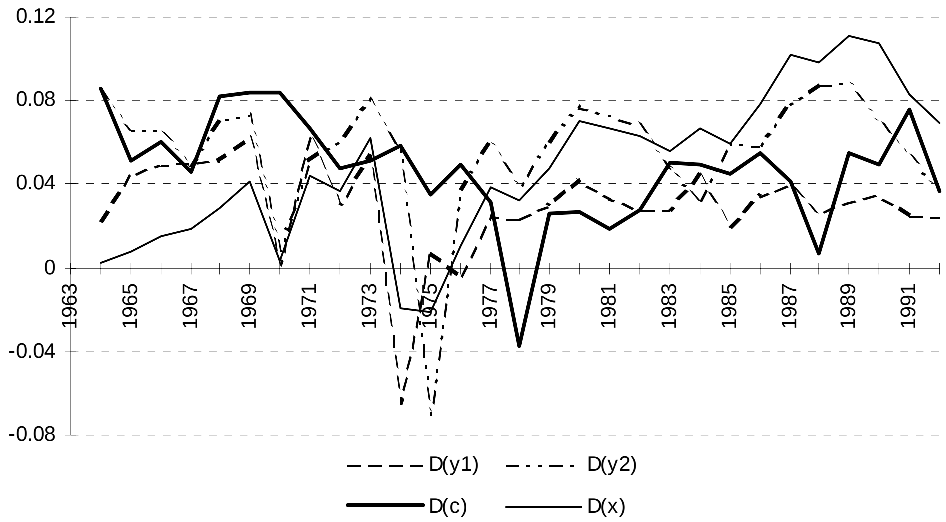

Figure 2: First differences of y_{1t}, y_{2t}, c and x.

order of the model. This considerably simplifies the analysis since in each step the test statistic follows the usual Dickey-Fuller distribution, and thus the well-known McKinnon critical values can be used.[18] In addition, we will also use a semiparametric analogue to the Dickey-Pantula procedure by sequentially employing the Phillips-Perron test.

It is commonly assumed that there are no economic time series being of higher order than $I(2)$. Therefore, in a first step, each series has to be tested whether it is $I(2)$, conditional on the assumption that it is at least $I(1)$. This, however, amounts to testing whether the first differences of the series are $I(1)$, which is a standard problem in unit root testing (see Haldrup 1998, 607). If the null hypothesis of non-stationarity can be rejected, the second step consists of testing $I(1)$ against $I(0)$ (stationarity). If the null cannot be rejected, however, we conclude that the series in fact has a double unit root. Table 1 reports the various test statistics for each of the series.

[18]See Haldrup (1998, section 5) for further details. This author also offers a survey and critique of various other tests for double unit roots.

Table 1: Tests for stationarity

variable	results					
	$ADF(1)$	$ADF(2)$	$ADF^+(1)$	$ADF^+(2)$	PP	PP^+
y_1	2.128	0.991	−2.158	−3.002	2.497	−1.121
y_2	2.210	1.957	−1.734	−1.277	3.963	−1.607
c	1.539	1.005	−1.890	−2.380	2.271	−2.194
x	1.729	1.492	−0.354	−0.160	8.044	0.338
Δy_1	−1.280	−1.421	−2.386	−2.574	-2.304^*	-4.970^{**}
Δy_2	−1.418	−1.026	-3.872^*	−3.040	−1.686	-4.133^*
Δc	−1.035	−0.876	−2.307	−1.845	−1.658	-3.672^*
Δx	−0.446	−0.315	−2.788	−2.838	0.066	−2.655

$ADF^+(j)$:= Augm. Dickey-Fuller (+: with constant and trend) with j lags
$PP^{(+)}$:= Phillips-Perron (+: with constant and trend)
$^*,^{**}$:= H_0 (nonstationarity) is rejected on the 5%, 1% significance level

First, consider the unit root tests on the levels of the variables. They all indicate that the initial assumption of the series being at least $I(1)$ is justifiable. Now, consider the first step in the Dickey-Pantula procedure, i.e. unit root tests on the first differences of each variable. Table 1 shows that the Augmented Dickey-Fuller (ADF) test and the Phillips-Perron (PP) test do not generally yield similar results. In particular, the PP^+ tests are somewhat at odds with the rest of the other results.[19] Nevertheless, the clear majority of the statistical tests confirms the visual impression of the existence of double unit roots in all of the series. In sum, the available evidence suggests that the variables should be taken as $I(2)$.[20]

It can be shown that $I(2)$ processes have a rather disturbing property which appears to make them impractical for the purposes of analyzing economic behavior: a shock that occurred in the past has an ever increasing effect on the series' level over time (see Granger 1997, 173). Haldrup (1998, 604), however, argues that such an effect might result from a shock to the growth rate of an already persistent time series. In the telecommunications

[19]This distortion might be generated by the fact that the relevant test statistic was calculated under the assumption of a constant and a trend. Recall, however, that all series were normalized at zero in the first observation.

[20]Interestingly, Hunt and Lynk (1990) find a similar result in their analysis of British Telecom data from 1951/52 to 1980/81.

industry considered here, Haldrup's argument appears perfectly reasonable, since the steady demand increase over time could well be interpreted to stem from such a growth shock in the past. Hunt and Lynk (1990, fn. 5) implicitly put forward a related argument. They explain their findings with the rapid expansion of telecommunications demand and argue that they are not implausible for microeconomic series in general. To be sure, there are further potential reasons for our finding of $I(2)$ variables, such as the small variance of the series most probably stemming from the aggregation of data with higher frequency[21] and the well-known sensibility of unit-root tests with respect to structural changes.[22]

4 Testing for Subadditivity

4.1 The Cost Function

As pointed out above, theoretical considerations call for a quadratic specification of the cost function. We use a somewhat adapted form of (2) in that we ignore the quadratic outputs terms y_{1t}^2 and y_{1t}^2,[23] and—following Baumol and Braunstein (1977)—allow the exponent γ of the interaction term $x_t = y_{1t}y_{2t}$ to be different from 1. The cost function to be estimated is thus

$$c_t = \beta_0 + \beta_1 y_{1t} + \beta_2 y_{2t} + \beta_3 x_t^\gamma. \tag{3}$$

It now has to be established theoretically for what parameter values of $\beta_0, \beta_1, \beta_2, \beta_3$ and γ this cost function exhibits subadditivity. Recall that a function is called locally subadditive if it is transray convex and features decreasing ray average costs. The transray convexity condition simply requires $\partial^2 c_t(\cdot)/\partial y_{1t}\partial y_{2t} = \beta_3\gamma^2 x_t^{\gamma-1} < 0$, that is $\beta_3 < 0$. The conditions for the existence of decreasing ray average costs are more subtle. First, set $y_{2t} = \kappa y_{1t}$,

[21] According to information by PTT, monthly or even daily observations do exist but are not made available to the general public.

[22] In fact, PTT's statistical yearbooks indicate that after 1974 the data for local and long distance calls is not comparable to earlier observations. The effects of this structural change on the estimations are discussed in more detail in section 4.

[23] This has the advantage of saving degrees of freedom.

Table 2: Parameter values for decreasing ray average costs

	value of the exponent γ	conditions for $d\left(\bar{c}_t^r\right)/dy_{1t} < 0$
case 1:	$1 \geq \gamma > \frac{1}{2}$	$\beta_0 > 0, \beta_3 < 0$
case 2:	$\gamma = \frac{1}{2}$	$\beta_0 > 0$
case 3:	$\frac{1}{2} > \gamma > 0$	$\left[\frac{\beta_0}{(2\gamma-1)\beta_3}\right]^{\frac{1}{\gamma}} > y_{1t}y_{2t}$

with $\kappa > 0$, to calculate costs along a ray through the origin, and divide by y_{1t} to obtain ray average costs

$$\bar{c}_t^r = \frac{\beta_0}{y_{1t}} + (\beta_1 + \beta_2\kappa) + \beta_3\kappa^\gamma y_{1t}^{(2\gamma-1)}. \tag{4}$$

Second, differentiate ray average costs $\bar{c}_t^r$ with respect to y_{1t} to obtain (see Baumol and Braunstein 1977, 1042)

$$\frac{d\left(\bar{c}_t^r\right)}{dy_{1t}} = -\frac{\beta_0}{y_{1t}^2} + (2\gamma-1)\beta_3\kappa^\gamma y_{1t}^{(2\gamma-2)}. \tag{5}$$

Decreasing ray average costs imply $d\left(\bar{c}_t^r\right)/dy_{1t} < 0$. Table 2 shows all three different cases of coefficient values for which this condition is satisfied.[24] We shall now compare the estimated coefficients of the cost function with the theoretically derived conditions for transray convexity and decreasing ray average costs and thereby try to determine whether the industry's cost function exhibits subadditivity or not.

[24]There is a little difference to the conditions reported by Baumol and Braunstein (1977, 1042) which, however, will not be of importance in the remainder of the paper.

Table 3: Coefficient Estimates

coefficients	specification (i)	specification (ii)
$\widehat{\beta}_0$	-0.038 (0.042)	-0.023 (0.030)
$\widehat{\beta}_1$	0.789 (0.656)	0.676 (0.269)
$\widehat{\beta}_2$	1.048 (0.566)	0.902 (0.288)
$\widehat{\beta}_3$	-0.794 (1.096)	-0.524 (0.185)
$\widehat{\gamma}$	0.825 (0.464)	1 ($-$)
number of observations	30	30
estimation method	NLLS	OLS
$\bar{R}^2$	0.978	0.979
SSR	0.078	0.079
DW	0.431	0.423
$ADF(1)$	-3.180	-3.071

$(\cdot)$:= Standard Error (Newey-West)
NLLS := Nonlinear Least Square
$\bar{R}^2$:= Adj. Coeff. of Determination
DW := Durbin-Watson
OLS := Ordinary Least Square
SSR := Sum of Square Residuals
$ADF(1)$:= Augm. Dickey-Fuller (Lag 1)

4.2 Estimation Results

Using the data set outlined in section 3, the estimation of equation (3) yields the results listed under specification (i) in Table 3. For the restricted version with $\gamma \equiv 1$, the results for specification (ii) apply. Figure 3 and 4 plot the estimated cost planes.

Observe that the exponent $\widehat{\gamma}$ lies within the range of case 1 in Table 2 for both specifications. Subadditivity therefore requires $\widehat{\beta}_0 > 0$ and $\widehat{\beta}_3 < 0$ in both cases. Inspection of Table 3 shows, however, that $\widehat{\beta}_0 < 0$ and $\widehat{\beta}_3 < 0$, independent of the specification of the cost function. Hence, we are tempted to conclude that our estimations do not indicate that the Swiss telecommunications industry was operating under the cost conditions of natural monopoly during the time observation.

Before drawing this conclusion, let us pause for a moment. Observe that the standard error of $\widehat{\beta}_0$ is larger than the value of the respective coefficient for both specifications. This is not much of a problem concerning the interpretation of the coefficients as long as they are estimated consistently. Both

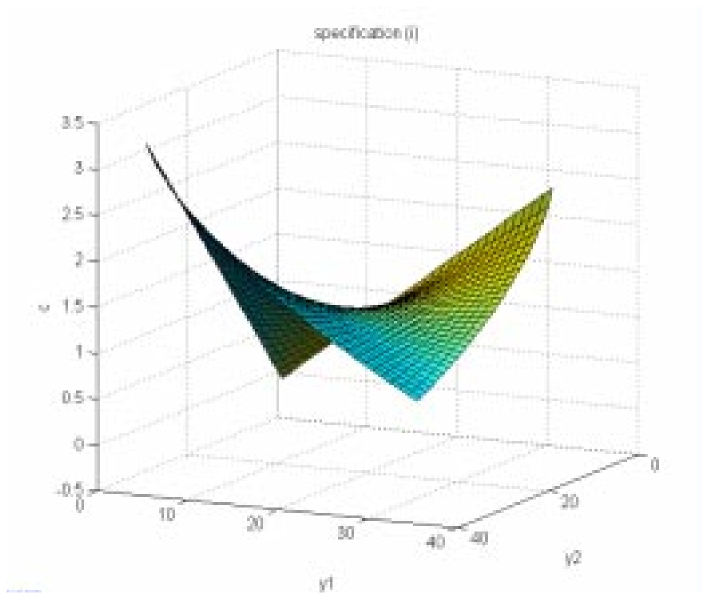

Figure 3: Estimated cost plane, specification (i)

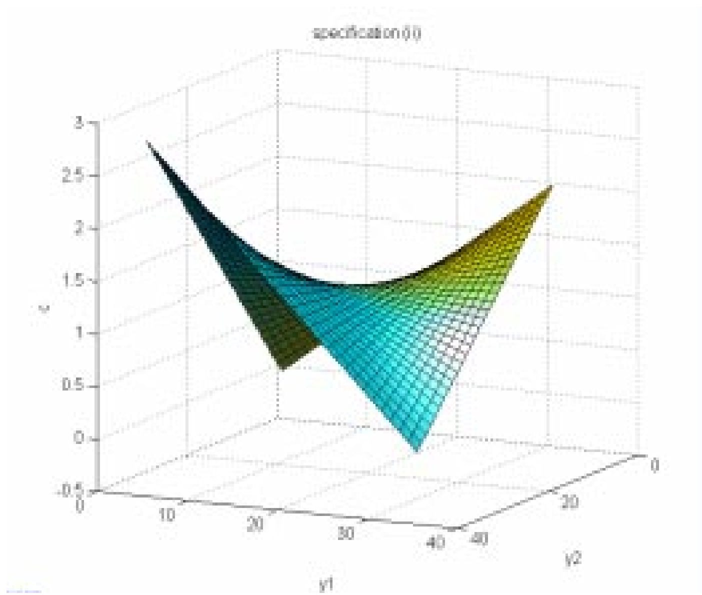

Figure 4: Estimated cost plane, specification (ii)

for $I(1)$ and $I(2)$ variables, however, (super)consistent coefficient estimates can only be attained if the series are *cointegrated*.[25] For the results to hold, it thus needs to be shown that the variables are in fact cointegrated.

Haldrup (1994, pp. 165) demonstrates how Engle and Granger's (1987) well-known 2-step procedure for testing for cointegration can be adapted to the case of $I(2)$ variables. We shall follow his approach and restrict our attention to the analysis of the OLS equation (i.e. specification (ii)).[26] The idea is to set up a cointegrating regression and test the residuals' integration order, using an ADF-test. As above, the null hypothesis H_0 holds that the residuals are non-stationary (i.e., $I(1)$ as opposed to $I(0)$). According to the test result given in Table 3, H_0 cannot be rejected at the 10% significance level, since the value of the $ADF(1)$ test statistic is larger than the critical value -3.90.[27] That is, we cannot reject that the residuals are non-stationary.

Since the residuals are non-stationary, the OLS equation has to be qualified as "spurious". This conclusion is supported by the relation between $\bar{R}^2$ and the DW-statistic. Granger and Newbold (1974) propose to call any regression "spurious" for which $\bar{R}^2 > DW$. This relation clearly applies to specification (ii), and it thus exhibits the well-known characteristics of a "spurious" regression: serial correlation in the residuals and non-degenerate limiting distributions of the estimated coefficients (see Banerjee et al. 1993; Phillips 1986).

4.3 Refinement Strategies

There are essentially three basic strategies to solve the specification problem detected above:

(i) remove irrelevant variables from the equation;

(ii) add relevant variables not taken into account and

---

[25] See Haldrup (1994). Recall that the components of a vector $\mathbf{x}_t$ are said to be cointegrated of order d, b, denoted by $\mathbf{x}_t \sim CI(d, b)$, if (i) all components of $\mathbf{x}_t$ are $I(d)$; (ii) there is a vector $\boldsymbol{\alpha}$ ($\neq 0$) so that $z_t = \boldsymbol{\alpha}'\mathbf{x}_t \sim I(d-b), b > 0$ (Engle and Granger 1987, 253).

[26] Since the coefficient estimates have the same signs and roughly similar values, we do not forego useful information by restricting ourselves to the analysis of specification (ii).

[27] Haldrup (1994, 168, Table 1) reports the adapted critical values for this ADF-test.

(iii) remove the non-stationarity in the data by differencing it sufficiently many times.

Unfortunately, none of these strategies helps to solve the problem in the case considered here, as it turns out. This is most obvious for strategy (i): the OLS equation is given by the economic theory laid out above and, in this most simple form, there is no irrelevant variable that could be removed from the regression. Adding another variable, as proposed by strategy (ii), needs to make economic sense as well. Therefore, we might add either price variables or—again for want of a more adequate measure—a linear time trend as a proxy-variable for the technological progress undoubtedly evolving in the telecommunications industry during the time of observation.[28] As mentioned above, price data are not available; further, adding a time trend does not solve the problem, either. On the contrary: the estimated equation does not even have the characteristics of a proper cost function anymore, since the marginal cost of producing output y_2 is now estimated to be negative ($\partial c_t(\cdot)/\partial y_{2t} < 0$). Hence, strategy (ii) is not helpful, either. There remains the possibility to further transform the data in order to remove the non-stationarity of the data (strategy (iii)). However, regressing the first or second differences on each other—all series in the equation are then $I(1)$ and $I(0)$, respectively—yields the following results: the residuals are still non-stationary and serially correlated and $\bar{R}^2$ proves to be negative.

A more elaborate form of implementing strategy (iii) consists of ironing out the structural change in the collection of the data mentioned above (see fn. 22). This structural change only applies to the collection of output data and might therefore disturb a potential long-run equilibrium relation between output and costs. In fact, a Chow-Test performed for 1974 rejects the null hypothesis of no structural change.[29] Interestingly, yet another way

[28]When working with time series data, the specification of technological progress can play an important role. With swift technological progress and increasing demand, there are two effects potentially leading to a decrease of unit costs: (i) technological progress that shifts down the average cost curve, and (ii) a growth of output that moves the firm down along its (given) average cost curve. These effects can only be disentangled if technological progress is adequately specified (see Panzar 1989, 54).

[29]The value of the test statistic is 22.29 and thus clearly larger than the critical value 2.82 from the F-distribution (see Greene 1993, 734).

of implementing strategy (iii) yields a somewhat different result. If the individual time series y_{1t} and y_{2t} are modeled as ARIMA-processes, a level shift is detected for y_{1t} (if only on the 10% significance level), but not for y_{2t}. However, if the adjusted data is—after ironing out the level shift and the remaining outliers—again used for the estimation of specification (ii), there are no changes to the coefficients worth mentioning.[30]

Hence, since all strategies available to refine the estimations of the cost function fail, we are left with the results summarized in Table 3. How are these results to be interpreted? In the case of stationary processes it is guaranteed that the coefficients of a regression on a set of independent variables converge to zero. This is not the case for integrated variables (Banerjee et al. 1993, p. 70). There is thus the possibility that the coefficients indicate a correlation even if there is none.[31] *In sum, based on the data publicly available we cannot determine with certainty whether the Swiss telecommunications industry qualified as a natural monopoly prior to regulatory reform.*

This ambiguous result is due to the fact that there does not seem to be a stable long-run relation between outputs and cost. There are two lines of arguments that explain this result, with the second being of both more theoretical and practical importance:

(i) The publicly available data is rather sparse. With 30 observations, the sample is small and the variance of the data appears to be extremely reduced as a consequence of the aggregation of higher frequency data. In fact, the data aggregation process might be an important reason for our finding of $I(2)$ variables.

(ii) There is a continuous change in the production conditions of dynamic industries, such as telecommunications, which is not properly reflected

[30] In addition, recursive estimations do not indicate an improved stability of the coefficient estimates.

[31] Multicollinearity is another potential issue. It refers to the problem that if the regressors are highly correlated, the coefficient estimates will not be precise. In particular, coefficients might have an implausible magnitute or even the wrong sign. Both effects are potentially relevant for our analysis, since the regressors y_{1t}, y_{2t} and x_t are correlated. Note, however, that multicollinearity is not a specification but a data problem for which no simple remedies exist; see Greene (1993, pp. 266) for further details.

in the conventional natural monopoly analysis. This analysis is based on the assumption of a static cost function and a given set of relevant products. Such an assumption is innocent for a short period of time, but hardly sustainable over a time stretch of several years. For instance, new and upcoming services like fax, data transmission, mobile telephony and internet access might have changed PTT's cost function in a way that is not reflected in the available data on the more traditional telecommunications services, often called "plain old telephone services" (POTS).

The second argument suggests that the static natural monopoly concept is ill suited as a theoretical concept for informing policy decisions. Since even if the production of a given set of products can be shown to be a natural monopoly at a given point in time, this does not imply that costs do not change, and that (potential) competitors cannot innovate, produce differentiated products, successfully compete and thereby generate welfare improvements.

5 Conclusions

The ambiguous result of the empirical analysis—which is in accord with the often contradicting results found in the literature—might not appear to be particularly helpful to regulators and policy makers. Note, however, that, as Berg and Tschirhart (1995, p. 121) point out,

> "in a world of uncertainties, little can be rigorously proven in the policy arena."

Our analysis demonstrates that it is very difficult to empirically substantiate the claim that the telecommunications industry—or any other industry—is a natural monopoly: during the time of observation considered in this study, there was a widespread consensus that Swiss telecommunications actually was a natural monopoly, and probably even few economists would have contested this view. Until now, however, we have not been able to come up with convincing empirical evidence to support it.

But even if a given industry actually was a (static) natural monopoly, policy makers and regulators alike should be cautious when being asked to protect it from 'inefficient' competition. It might well be that such protection prevents innovation that eventually transforms the incumbent to an 'unnatural' monopoly. Swiss telecommunications is a case in point.

References

Banerjee, A., Dolado, J.J., Galbraith, J.W., Hendry, D.F. (1993): *Co-Integration, Error Correction, and the Econometric Analysis of Non-Stationary Data.* Oxford.

Baumol, W.J. (1977): "On the Proper Cost Test for Natural Monopoly in a Multiproduct Industry", *American Economic Review*, Vol. 67, No. 5, 809-822.

Baumol, W.J., Braunstein, Y.M. (1977): "Empirical Study of Scale Economies and Production Complementarity: The Case of Journal Publication", *Journal of Political Economy*, Vol. 85, No. 5, 1037-1048.

Baumol, W.J., Panzar, J.C. , Willig, R.D. (1982): *Contestable Markets and the Theory of Industry Structure.* New York.

Berg, S.V., Tschirhart, J. (1995): "A Market Test for Natural Monopoly in Local Exchange", *Journal of Regulatory Economics*, Vol. 8, No. 2, 103-124.

Blackman, C.R. (1995): "Universal Service: Obligation or Opportunity", *Telecommunications Policy*, Vol. 19, No. 3, 171-176.

Bond, J. (1997): "Telecommunications is Dead. Long Live Networking", in: World Bank (Ed.): *The Information Revolution and the Future of Telecommunications*, June 1997, 11-14.

Bräutigam, R.R. (1979): "Optimal Pricing with Intermodal Competition", *American Economic Review*, Vol. 69, No. 1, 38-49.

Buehler, S. (1999): *A Further Look at Two-way Network Competition in Telecommunications*, Working Paper 9904 of the Socioeconomic Institute (revised version), University of Zurich.

Buehler, S. (1998): "Regulatory Reform of Telecommunications in Switzerland", *Telecommunications Policy*, Vol. 22, No. 8, 671-680.

Charnes, A., Cooper, W.W., Sueyoshi, T. (1988): "A Goal Programming/Constrained Regression Review of The Bell System Breakup", *Management Science*, Vol. 34, 1-26.

Engle, R.F., Granger, C.W.J. (1987): "Co-Integration and Error Correction: Representation, Estimation and Testing", *Econometrica*, Vol. 55, No. 2, 251-276.

Evans, D.S., Heckman, J.J. (1984): "A Test for Subadditivity of the Cost Function with an Application to the Bell System", *American Economic Review*, Vol. 74, No. 4, 615-623.

Evans, D.S., Heckman, J.J. (1983): "Multiproduct Cost Function Estimates and Natural Monopoly Tests for the Bell System" in: Evans, D.S. (ed.): *Breaking Up Bell*. Amsterdam.

Granger, C.W.J. (1997): "On Modelling the Long Run in Applied Economics", *Economic Journal*, Vol. 107, No. 1, 169-177.

Granger, C.W.J, Newbold, P. (1974): "Spurious Regression in Econometrics", *Journal of Econometrics*, Vol. 2, No. 2, 111-120.

Greene, W.H. (1993): *Econometric Analysis* (third edition). New Jersey.

Haldrup, N. (1998): "An Econometric Analysis of I(2) Variables", *Journal of Economic Surveys*, Vol. 12, No. 5, 595-650.

Haldrup, N. (1994): "The Asymptotics of Single-Equation Cointegration Regressions with I(1) and I(2) Variables, *Journal of Econometrics*, Vol. 63, No. 1, 153-181.

Hamilton, J.D. (1994): *Time Series Analysis*. Princeton, New Jersey.

Hunt, L.C., Lynk, E.L. (1990): "Divesture of Telecommunications in the UK: A Time Series Analysis", *Oxford Bulletin of Economics and Statistics*, Vol. 52, No. 3, 229-251.

Jamison, M. (1998): *A Further Look at Proper Cost Tests for Natural Monopoly*. University of Florida, Gainesville (mimeo).

Knieps, G. (1997): "Phasing out Sector-Specific Regulation in Competitive Telecommunications", *Kyklos*, Vol. 50, No. 3, 325-339.

Laffont, J.J., Tirole, J. (1996): "Creating Competition Through Interconnection: Theory and Practice", *Journal of Regulatory Economics*, Vol. 10, No. 3, 227-256.

Manzini, A., Thalmann, P. (1994): "The Dynamic Allocative Efficiency of a Public Utility: Swiss Telecommunications", *Swiss Journal of Economics and Statistics*, Vol. 130, No. 2, 129-144.

Mitchell, B., Neu, W., Neumann, K.H., Vogelsang, I. (1995): "The Regulation of Pricing of Interconnection Services", in: Brock, G. (ed.): *Toward a Competitive Telecommunication Industry. Selected Papers from the 1994 Telecommunications Policy Research Conference.* New Jersey, 95-118.

Noam, E.M. (1994): "Beyond Liberalization III: Reforming Universal Service", *Telecommunications Policy*, Vol. 18, No. 9, 687-704.

Panzar, J.C. (1989): "Technological Determinants of Firm and Industry Structure", in: Schmalensee, R., Willig, R. (Eds.): Handbook of Industrial Organization, Vol. 1. Amsterdam, 3-62.

Phillips, P.C.B. (1986): "Understanding Spurious Regresssions in Econometrics", *Journal of Econometrics*, Vol. 33, No. 3, 311-340.

Röller, L.H. (1990a): "Proper Quadratic Cost Functions with an Application to the Bell System", *Review of Economics and Statistics*, Vol. 72, No. 2, 202-210.

Röller, L.H. (1990b): "Modelling Cost Structure: the Bell System Revisited", *Applied Economics*, Vol. 22, No. 12, 1661-1674.

Shin, R.T., Ying, J.S. (1992): "Unnatural Monopolies in Local Telephone", *Rand Journal of Economics*, Vol. 23, Nr. 2, 171-183.